

Conformational Search Using a Molecular Dynamics–Minimization Procedure: Applications to Clusters of Coulombic Charges, Lennard–Jones Particles, and Waters

ZHENQIN LI, KEITH E. LAIDIG, VALERIE DAGGETT

Department of Medicinal Chemistry, University of Washington, Seattle, Washington 98195-7610

Received 3 February 1997; accepted 12 August 1997

ABSTRACT: A new conformational search method, molecular dynamics–minimization (MDM), is proposed, which combines a molecular dynamics sampling strategy with energy minimizations in the search for low-energy molecular structures. This new method is applied to the search for low energy configurations of clusters of coulombic charges on a unit sphere, Lennard–Jones clusters, and water clusters. The MDM method is shown to be efficient in finding the lowest energy configurations of these clusters. A closer comparison of MDM with alternative conformational search methods on Lennard–Jones clusters shows that, although MDM is not as efficient as the Monte Carlo–minimization method in locating the global energy minima, it is more efficient than the diffusion equation method and the method of minimization from randomly generated structures. Given the versatility of the molecular dynamics sampling strategy in comparison to Monte Carlo in treating molecular complexes or molecules in explicit solution, one anticipates that the MDM method could be profitably applied to conformational search problems where the number of degrees of freedom is much greater. © 1998 John Wiley & Sons, Inc. *J Comput Chem* 19: 60–70, 1998

Keywords: molecular-dynamics minimization; conformational search; water cluster; Lennard–Jones clusters; Coulombic clusters

Correspondence to: V. Daggett; e-mail: daggett@u.washington.edu

Contract/grant sponsor: Office of Naval Research; contract/grant number: N0014-95-0484

Contract/grant sponsor: American Health Assistance Foundation

Introduction

The problem of predicting the most energetically stable structures of large molecular systems is notoriously difficult. Even if the precise potential energy function or potential of mean force describing the atomic interactions of a molecular system is known, the process of finding the global minimum involves the search of a conformational space whose number of alternative potential minima increases exponentially with the number of degrees of freedom. The multiple-minimum problem encountered during global minimization^{1,2} is, in general, a computationally difficult problem and is known in computer sciences terminology as an NP-hard problem.^{3–5}

Over the past decade, many computational algorithms and procedures have been proposed that, coupled with recent advances in computational power of computer hardware, have achieved varying degrees of success using complete potential energy functions in the global minimization of systems such as organic molecules,^{6–14} peptides or proteins (or parts of proteins),^{4,8,15–34} and clusters of Lennard–Jones atoms^{35–44} and water molecules.^{43,45–50} However, aside from the problems of improving the accuracy of potential energy functions, potentials of mean force, or other empirical objective functions, the efficiency of the conformational search is still the bottleneck in the applications of many structure-prediction and simulation methods.

In this article, we describe a new conformational search method, the molecule dynamics–minimization (MDM) procedure, which is analogous to the Monte Carlo–minimization (MCM) algorithm proposed by Li and Scheraga.⁴ (Both MCM and MDM satisfy statistical mechanics in the sense that the conformational states of energy minima conform to a Boltzmann distribution. The original MCM procedure samples the conformational superlattice of energy minima via a metropolis Monte Carlo algorithm. In contrast, this new MDM procedure “searches” the conformational superlattice using an MD-like sampling strategy, utilizing the additional information about the potential energy surface contained within the MD velocities. This approach is analogous to the “smart (or biased) random move” in Monte Carlo sampling,^{16,22} but without compromising the Boltzmann statistics obeyed by the energy minima.

Methods

CONCEPTUAL FRAMEWORK

As in the Monte Carlo–minimization procedure,⁴ we describe the multidimensional conformational space in terms of a reduced (“coarse-grained”) representation of conformational states such that all structures in the neighborhood of a single potential well are represented by a single state (i.e., the minimum-energy configuration). Unlike MCM, however, the transition to a new conformational state is driven by a dynamic process analogous to Newton’s second law, instead of random sampling as in the Monte Carlo procedure. The algorithm initiates its move from the present position by stepping along the MD velocities. Because the velocities result, to some extent, from the forces within the system, as well as the kinetic energy, these steps are governed by the potential energy surface. This “direction” of the sampling should avoid many configurations of high steric energy that would be encountered using a random Monte Carlo strategy. To ensure Boltzmann statistics for the ensemble, we enforce conservation of the total energy during the transition between two energy minima. The accelerations necessary to allow a quadratic step between the two minima are determined, assuming a valid step can be found in this “pseudo-Newtonian” step. The corresponding velocities, and the resulting kinetic energy, are adjusted to provide conservation of the total energy. Analogous to constant temperature MD,^{51,52} the velocities are randomly updated according to a Maxwellian distribution appropriate for the temperature of interest to ensure Boltzmann statistics. Thus, the collection of transitions does not provide a continuous trajectory as in classical MD methods.

IMPLEMENTATION

The algorithm for molecular dynamics–minimization is an iterative procedure.

Step 1. Generate new conformation. Starting with the current conformation, \mathbf{x}_i^0 , and velocities, \mathbf{v}_i^0 , we construct a plausible new conformation:

$$\mathbf{x}'_i = \mathbf{x}_i^0 + \mathbf{v}_i^0 \tau \quad (1)$$

where τ is, in essence, a kinetic control parameter on the time scale of thermal collisions, which is on the picosecond time scale at room temperature.

This step differs from the random sampling step in the MCM procedure in that the displacements are directed along the current velocities.

Step 2. Perform conventional energy minimization (either full or partial). The goal of this step is to remove any steric clashes created within the new conformation obtained in step 1. This also serves to reduce the representation of the conformational states from that of a continuous set of point states to a discrete set of energy minima:

$$\mathbf{x}_i' \rightarrow \mathbf{x}_i^{\text{new}} \quad (2)$$

where $\mathbf{x}_i^{\text{new}}$ is the location of the new minimum obtained by energy minimization from the position \mathbf{x}_i' .

Step 3. Ensure energy conservation. The Boltzmann statistics for the conformational energy minima can be ensured if we impose conservation of total energy during the transition from \mathbf{x}_i^0 to $\mathbf{x}_i^{\text{new}}$, while allowing the velocities to be updated incrementally, afterward, according to a Maxwellian distribution (see step 4). However, while MD connects two distinct states along the trajectory by many iterations of a finite-difference implementation of Newton's equation at small time intervals, we try to connect two distinct energy minima via a single pseudo-Newtonian iteration such that:

$$\mathbf{v}_i^{\text{new}} = \mathbf{v}_i^0 + \mathbf{a}_i \tau \quad (3)$$

$$\mathbf{x}_i^{\text{new}} = \mathbf{x}_i^0 + \mathbf{v}_i^0 \tau + \mathbf{a}_i \tau^2 / 2 \quad (4)$$

where \mathbf{a}_i is the effective acceleration of atom i caused by an effective force along the trajectory between the two minima \mathbf{x}_i^0 and $\mathbf{x}_i^{\text{new}}$, τ is the effective time interval needed for the conformational transition from \mathbf{x}_i^0 to $\mathbf{x}_i^{\text{new}}$ to take place.

For the total energy to be conserved, the new positions $\mathbf{x}_i^{\text{new}}$ and velocities $\mathbf{v}_i^{\text{new}}$ need to satisfy:

$$\begin{aligned} U(\mathbf{x}_i^{\text{new}}) + \Sigma 1/2 m_i (\mathbf{v}_i^0 + \mathbf{a}_i \tau)^2 \\ = U(\mathbf{x}_i^0) + \Sigma 1/2 m_i (\mathbf{v}_i^0)^2 \end{aligned} \quad (5)$$

Unlike conventional molecular dynamics, where the acceleration \mathbf{a}_i is known directly from the potential energy function $U(\mathbf{x}_i)$ via Newton's second law, $\mathbf{a}_i = -\nabla U(\mathbf{x}_i)/m_i$, here the effective acceleration \mathbf{a}_i connecting the two distinct states is not known. However, if we know the new position $\mathbf{x}_i^{\text{new}}$, we can use eqs. (4) and (5) to indirectly deduce the values of \mathbf{a}_i and τ , provided such

values exist. Substituting from eq. (4):

$$\mathbf{a}_i \tau = 2[(\mathbf{x}_i^{\text{new}} - \mathbf{x}_i^0)/\tau - \mathbf{v}_i^0] \quad (6)$$

into eq. (5), we have

$$\begin{aligned} U(\mathbf{x}_i^{\text{new}}) + \Sigma 1/2 m_i [2(\mathbf{x}_i^{\text{new}} - \mathbf{x}_i^0)/\tau - \mathbf{v}_i^0]^2 \\ = U(\mathbf{x}_i^0) + \Sigma 1/2 m_i (\mathbf{v}_i^0)^2 \end{aligned}$$

or:

$$\Sigma m_i (\Delta \mathbf{x}_i)^2 / \tau^2 - \Sigma m_i \Delta \mathbf{x}_i \cdot \mathbf{v}_i^0 / \tau + \Delta U / 2 = 0 \quad (7)$$

where $\Delta \mathbf{x}_i = \mathbf{x}_i^{\text{new}} - \mathbf{x}_i^0$ and $\Delta U = U(\mathbf{x}_i^{\text{new}}) - U(\mathbf{x}_i^0)$. Eq. (7) is quadratic in $1/\tau$, whose meaningful solution requires that

$$(\Sigma m_i \Delta \mathbf{x}_i \cdot \mathbf{v}_i^0)^2 - 2 \Delta U \left[(\Sigma m_i (\Delta \mathbf{x}_i)^2) \right] \geq 0 \quad (8)$$

If the condition in eq. (8) is not satisfied, it means the step $\Delta \mathbf{x}_i = \mathbf{x}_i^{\text{new}} - \mathbf{x}_i^0$ is too large for the existence of real values of \mathbf{a}_i and τ to satisfy both the pseudo-Newtonian equation, [eq. (4)], and the energy conservation condition [eq. (5)]. This move, therefore, must be aborted and we return to step 1 to try a smaller value of τ . If the inequality condition in eq. (8) holds, there are usually two possible real solutions for $1/\tau$:

$$\begin{aligned} 1/\tau = \left\{ \Sigma m_i \Delta \mathbf{x}_i \cdot \mathbf{v}_i^0 \pm \text{sqrt} \left[(\Sigma m_i \Delta \mathbf{x}_i \cdot \mathbf{v}_i^0)^2 \right. \right. \\ \left. \left. - 2 \Delta U \Sigma m_i (\Delta \mathbf{x}_i)^2 \right] \right\} / \left[\Sigma 2 m_i (\Delta \mathbf{x}_i)^2 \right] \end{aligned} \quad (9)$$

We choose the positive branch, because it corresponds to the typical Newtonian move along a constant potential surface when $\Delta U = 0$ and $\Delta \mathbf{x}_i = \mathbf{v}_i^0 \tau$. Using the effective time interval calculated in eq. (9), we can obtain the effective acceleration in eq. (6) and, from that, the new velocity in eq. (3).

Step 4. Update velocities from the Maxwellian distribution. To ensure the Boltzmann statistics for the conformational energy minima, we need not only conservation of total energy during the transition from \mathbf{x}_i^0 to $\mathbf{x}_i^{\text{new}}$ (step 3) but also to update the velocities according to the appropriate Maxwellian distribution. This is similar to the typical practice in constant temperature molecular dynamic (MD) simulations,^{51,52} where the velocities are scaled periodically. In our algorithm, after each step 3 we randomly select a number of atoms (with an average of two) to have their velocities updated from the Maxwellian distribution; the velocities of all

other atoms in the system are unaffected. After step 4 is completed, we go back to step 1 with the updated positions and velocities, $\mathbf{x}_i^{\text{new}} \rightarrow \mathbf{x}_i^0$ and $\mathbf{v}_i^{\text{new}} \rightarrow \mathbf{v}_i^0$, and the iteration process is repeated. The total number of MDM iterations that satisfy the inequality in eq. (8) divided by the total number of trial minimizations give a measure of the acceptance ratio, which is conceptually similar to that used in Metropolis Monte Carlo simulations. In our computations, we usually choose our step size, τ , such that the acceptance ratio is between 40% and 60%.

Results

THE THOMSON PROBLEM

As an initial test of the conformational searching of the molecular dynamics–minimization procedure, we investigated the Thomson problem.⁵³ Given N point charges on the surface of a unit conducting sphere interacting only through their mutual coulombic forces, what is the configuration of the charges for which the Coulomb energy $\sum_{i < j} 1/|\mathbf{r}_i - \mathbf{r}_j|$ is minimized? This question was originally asked by Thomson⁵³ for $2 \leq N \leq 100$, and has recently been investigated by many investigators.^{54–60} Contrary to intuition, the configuration of minimum energy is not necessarily the one that maximizes the distance between each pair or the configuration of greatest symmetry. Erber and Hockney⁵⁵ compiled a table of the lowest energy configurations as well as some of the higher energy ones for $2 \leq N \leq 65$ by minimization of the Coulomb energy of randomly generated configurations. Altschuler et al.⁵⁸ published a table of lowest energy configurations for $66 \leq N \leq 100$ obtained by a method of “constrained global optimization”; three of these were later corrected by Erber and Hockney⁵⁹ as well as by Altschuler et al.⁶⁰

We have carried out MDM iterations from random starting structures to find lowest energy configurations for coulombic charges on a unit sphere, with $66 \leq N \leq 100$. The resulting lowest Coulomb energies are listed in Table I, and they confirm the result of Edmunsen⁵⁶ for $N = 78$ (which was found based on symmetry); those of Erber and Hockney⁵⁹ for $N = 69, 86$, and 87 ; and those of Altschuler et al.⁵⁸ for the other energies between $66 \leq N \leq 100$. In most of these simulations, the lowest energy configuration for each system can be found in no more than five MDM iterations.

As pointed out by Erber and Hockney,⁵⁹ based on their computer trials in the range of $70 \leq N \leq 112$, the number of distinct configurations associated with each value of N grows exponentially; that is:

$$M(N) \approx 0.382 \exp(0.0497N) \quad (10)$$

The identification of global minima among such a large set of nearly degenerate states for a complex system of this type poses a formidable technical challenge. Nonetheless, for this test case, the MDM procedure is efficient and generic and was able to locate the global minimum configuration for $N = 78$, which was not found by the “constrained global optimization” method of Altschuler et al.⁵⁸ (see Table I, footnote c)

TABLE I.
Lowest Coulombic Energies for Thomson’s Problem of N Unit Charges on the Surface of a Unit Sphere Obtained Using the Molecular Dynamics–Minimization (MDM) Procedure^a.

| N | Coulombic energy | N | Coulombic energy |
|-----|---------------------------|-----|---------------------------|
| 66 | 1882.441 525 | 84 | 3103.465 124 |
| 67 | 1942.122 700 | 85 | 3180.361 443 |
| 68 | 2002.874 702 | 86 | 3258.211 606 ^b |
| 69 | 2064.533 483 ^b | 87 | 3337.000 750 ^b |
| 70 | 2127.100 902 | 88 | 3416.720 197 |
| 71 | 2190.649 906 | 89 | 3497.439 019 |
| 72 | 2255.001 191 | 90 | 3579.091 223 |
| 73 | 2320.633 884 | 91 | 3661.713 699 |
| 74 | 2387.072 982 | 92 | 3745.291 636 |
| 75 | 2454.369 689 | 93 | 3829.844 338 |
| 76 | 2522.674 872 | 94 | 3915.309 270 |
| 77 | 2591.850 152 | 95 | 4001.771 676 |
| 78 | 2662.046 475 ^c | 96 | 4089.154 010 |
| 79 | 2733.248 357 ^d | 97 | 4177.533 600 |
| 80 | 2805.355 876 | 98 | 4266.822 464 |
| 81 | 2878.522 830 | 99 | 4357.139 163 |
| 82 | 2952.569 675 | 100 | 4448.350 634 |
| 83 | 3027.528 489 | | |

^a These results, believed to be the global energy minima, were first obtained by Altschuler et al.,⁵⁸ using a method of “constrained global optimization,” unless otherwise indicated.

^b Ref. 59, obtained based on minimization of randomly generated configurations.

^c Ref 56, obtained based on symmetry. Altschuler et al.⁵⁸ obtained 2622.047 213.

^d Altschuler et al.⁵⁸ obtained 2733.248 358.

LENNARD-JONES CLUSTERS

Unlike the Thomson problem, Lennard-Jones clusters provide a simple model problem relevant to the study of real physical systems such as rare gas clusters. The problem of finding the ground-state structures formed by clusters of atoms interacting through the pairwise Lennard-Jones potential has been comprehensively reviewed by Hoare.³⁵ Based on exhaustive enumeration of minimum energy structures for Lennard-Jones clusters with $N \leq 13$, it is shown that the total number of minimal energy structures scales exponentially with N as in the following expression^{35,61}:

$$M(N) \approx 2.444 \times 10^{-3} \exp(0.9931N) \quad (11)$$

We can see that the total number of distinct minima increases much faster with large N for the Lennard-Jones clusters than for the clusters of Coulomb charges, eq. (10). This difference can be explained in part by the general observation that the longer the range of the pairwise potential, the simpler the resulting multidimensional potential surface.^{62,63} In addition, other factors that contribute to the complexity of the potential energy surface of the Lennard-Jones system include the presence of a cluster boundary and the three-dimensional nature of the systems, whereas the Coulomb charges are limited to the surface of a unit sphere.

If the trend illustrated in eq. (11) continues for large N , it is not practical to perform an undirected search for all the local minima of the Lennard-Jones potential surface to find the global minimum. Current efforts in the search of the global minimum of Lennard-Jones clusters fall into two broad categories: (i) algorithms such as sphere packing, growth sequences, and (icosahedrally derived) lattice-based search and optimization,³⁸ which utilize physical intuition about the nature and symmetry of the Lennard-Jones system in its stable states; and (ii) the application of generic conformational search techniques such as molecular dynamics,³⁶ Monte Carlo,^{37,41} simulated annealing,³⁹ the diffusion equation method,⁴⁰ Gaussian phase packet/Gaussian density annealing,^{42,43} and the genetic algorithm.⁴⁴

At present, the lattice-based search and optimization method of Northby,³⁸ which is tailored to the study of atomic clusters (including the Lennard-Jones clusters), provides the most comprehensive results on the stable (presumed global

minimum⁷⁵) structures of Lennard-Jones clusters for $N \leq 147$. These results serve as a benchmark for generic conformational search techniques.

We consider our molecular dynamics-minimization method to be in the same category as the other generic conformational search techniques applied to the Lennard-Jones clusters in which the full potential energy surface is used. No prior knowledge about the Lennard-Jones clusters is utilized in our MDM application, except for the pairwise Lennard-Jones potential in reduced units:

$$V_{\text{LJ}}(r) = \sum_{i < j} (r_{ij}^{-12} - 2r_{ij}^{-6}) \quad (12)$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the interparticle distance.

MDM is able to locate the global minimum of a 56-membered Lennard-Jones cluster, the largest cluster successfully minimized by any of the methods using a full description of the potential energy hypersurface. Pillardy and Piela have reported successfully finding minima of clusters of up to 66 particles in size using a deformed hypersurface approach.⁷⁵ But, the lack of an obvious manner in which to deform the hypersurface of such complicated systems as explicitly solvated proteins has limited our interest in this type of algorithm. For more conventional methods, the diffusion equation method⁴⁰ found the minimum of a 55-particle cluster. But, MDM is able to handle significantly larger systems than the other approaches used in molecular studies. For example, approaches such as minimization of random structures,⁴¹ simulated annealing,³⁹ or Gaussian-phase-packet/Gaussian-density annealing algorithms⁴³ found minima for 28-, 25-, and 19-membered clusters, respectively. It is likely that these methods are not limited by the published cluster sizes, as MDM is not limited by the 56-particle cluster. Our intent here is to demonstrate that MDM is as capable in global searching as other methods, although application of MDM to more complicated systems is our ultimate goal.

WATER CLUSTERS

The potential energy surface of the Lennard-Jones clusters is more complex than that of coulombic charges on a unit sphere. However, the isotropic interactions of a Lennard-Jones cluster are much more simple than those in a cluster of water molecules, which introduces the additional complexity of the intramolecular degrees of freedom. As a result, methods such as the lattice-based search and optimization,³⁸ which work beautifully

for Lennard–Jones clusters, are no longer applicable to water clusters. Therefore, we have to rely on the more generic conformational search methods to investigate the structures and stability of more complicated molecular systems, such as water clusters.

Given the ubiquitous importance of water as a molecular liquid, the study of water clusters has important implications in untangling the molecular details, of, for example, aqueous solvation, cooperatively in hydrogen bonding, and the role of bound waters to macromolecules (Table II). As such, water clusters have been extensively investigated both experimentally and theoretically (reviewed in ref. 64a). A variety of conformational search methods have been applied to the study of the ground-state structures of small water clusters^{45–50, 65} using different rigid-body potential energy functions. In addition, high-level *ab initio* calculations have been carried out for some of these structures.^{68, 69}

To further test our MDM conformational search method, we address water clusters using the flexible-geometry three-centered water potential, F3C.⁷³ To our knowledge, this is the first systematic study of water clusters using a flexible-geometry, rather than a rigid-body, force field. Although the use of a flexible-geometric potential increases the number of degrees of freedom in the system, it has an advantage in molecular dynamics because, unlike a rigid-body force field for water, the water's degrees of freedom can be treated in the same way as those of other parts of the system.

The results of our MDM conformational search on water clusters (of up to nine water molecules) are summarized in Table III. The global minimal energy structures of the water trimer, tetramer, and pentamer are displayed in Figure 1. The geometries of the clusters are in reasonable agreement with the recent experimental investigations of Saykally et al.⁶⁴ The average O—O distance of 2.81 Å in all of the F3C-predicted clusters com-

TABLE II. Previous Conformational Search Studies on Water Clusters.

| Number of waters | Method of study ^a | Force field(s) ^b | Notes |
|------------------|--|--|----------------|
| 2 | Minimization ⁶⁵ DEM ⁴⁷ | Refs. 66–68 MCY | |
| 3 | Minimization ⁶⁵ DEM ⁴⁷ MD ⁴⁹ | Refs. 66–68 MCY CKL | |
| 4 | Minimization ⁶⁵ Monte Carlo ⁴⁵ DEM ⁴⁷ MD ⁴⁹ | Refs. 66–68 MCY MCY CKL | |
| 5 | Minimization ⁶⁵ Monte Carlo ⁴⁵ DEM ⁴⁷ MD ⁴⁹ | Refs. 66–68 MCY MCY CKL | |
| 6 | Monte Carlo ⁴⁵ DEM ⁴⁷ Minimization ⁴⁸ MD ⁴⁹ | MCY MCY MMC, MCY, OPLS, TIP4P CKL | |
| 7 | Monte Carlo ⁴⁵ DEM ⁴⁷ | MCY MCY | |
| 8 | Monte Carlo ⁴⁵ Monte Carlo ⁴⁶ DEM ⁴⁷ MD ⁴⁹ | MCY TIP3P MCY CKL | GM not reached |

^a DEM: diffusion equation method⁴⁷; MD: molecular dynamics.

^b Acronyms used for the potential energy functions for water: MCY stands for the potentials used in ref. 69; CPK stands for the potentials used in ref. 70; MMC stands for molecular mechanics for clusters from Ref. 71; TIP3P, TIP4P and OPLS (optimized potentials for liquid simulation) are the potentials developed by the Jorgensen group.⁷²

TABLE III.
Lowest-Energy Structures of Water Clusters
Obtained Using Molecular Dynamics–Minimization
(MDM) Conformational Search with F3C Potential.

| Number of waters | Lowest energy (kcal/mol) | Excitation energy ^a (kcal/mol) | Binding energy ^b (kcal/mol) |
|------------------|--------------------------|---|--|
| 2 | −6.703 | 0.072 | 6.703 |
| 3 | −19.402 | 5.872 | 12.699 |
| 4 | −31.943 | 6.159 | 12.541 |
| 5 | −42.026 | 1.402 | 10.083 |
| 6 | −52.353 | 0.576 | 10.327 |
| 7 | −63.761 | 0.050 | 11.408 |
| 8 | −78.622 | 0.306 | 14.861 |
| 9 | −90.228 | 1.020 | 11.606 |

^a Excitation energy is defined to be the energy difference between the global minimum energy and the next-lowest energy of an alternative local minimum.

^b Binding energy is defined to be $\Delta E_b = E_N - E_{N-1}$.

pares well with the experimental values of 2.98 Å, 2.90 Å, and 2.76 Å in the trimer, tetramer, and pentamer, respectively.^{64b} The predicted minima for the trimer and tetramer are planar, whereas the best experimental models show small out-of-plane rotations of the water molecules. But, high-level *ab initio* calculations of the “flipping” of the water

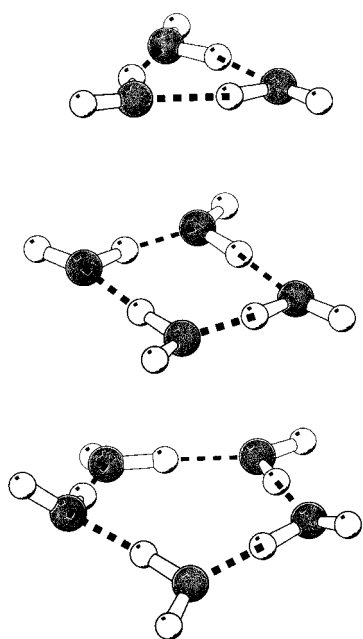


FIGURE 1. Lowest energy structures of water clusters obtained using the MDM procedure with the F3C potential. Top: trimer; middle: tetramer; and bottom: pentamer.

molecules through the plane of the trimer⁷⁴ find a barrier of only 0.26 kcal mol^{−1}, suggesting that the planar predictions of our simple water model are reasonable. The MDM-predicted pentamer shows the small out-of-plane rotations of waters predicted by the best experimental models.^{64c} The experimental characterization of the water hexamer is less clear,^{64d} with a cage structure being in best agreement with the vibrational–rotational tunneling spectra. High-level theory finds a number of low-lying, nearly degenerate hexamer structures.^{46b} After correction for zero-point energy differences, calculations predict that the cage structure is most stable, followed by a prism structure 0.2 kcal mol^{−1} less stable and then by an “open-book” structure about 3 kcal mol^{−1} higher in energy.^{46b} The relative energies of the cage and prism are reversed without quantum vibrational corrections. The lowest energy F3C hexamer is the book structure (Fig. 2a) with a cage structure (Fig. 2b) being 2 kcal mol^{−1} less stable. MDM did not find a corresponding prism structure. The discrepancies between the F3C results and the experiment and high-level theory on the hexamers are likely due to the lack of quantum vibrational corrections and quantum mechanical many-body effects in our classical water model. For octamer clusters,^{69c} unlike the *ab initio* calculations, MDM found a D₂ structure (Fig. 3a) about 0.3 kcal mol^{−1} more sta-

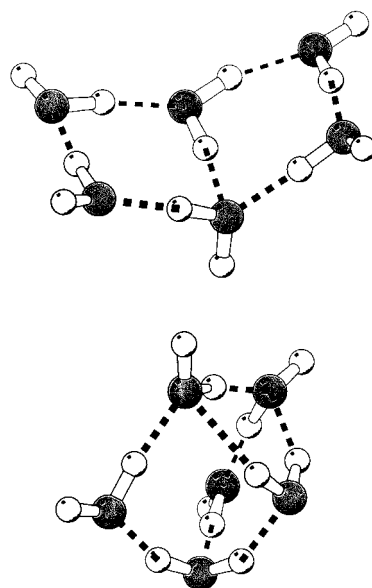


FIGURE 2. Lowest energy hexamer water clusters. (a) Book structure. (b) Cage structure.

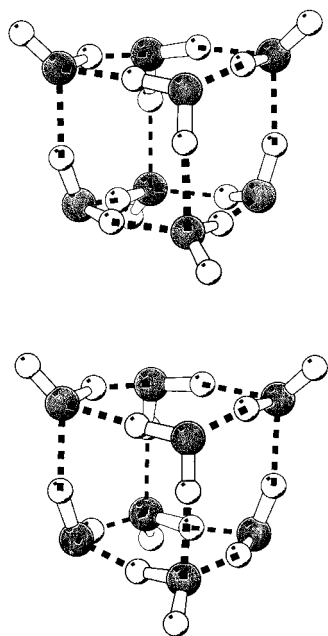


FIGURE 3. Lowest energy octamer water clusters. (a) D_2 structure. (b) S_4 structure.

ble than a S_4 structure (Fig. 3b), but MDM did not find the C_1 structure proposed by Stillinger et al.^{67b}

Discussion

Although many conformational search methods have been developed and applied to different molecular systems, there have been few systematic comparisons of the different methods. Part of the difficulty of making such a comparison is that many methods are tailored to specific systems that are not easily comparable. In the case of studies of water clusters, where the same system has been studied using different conformational search methods, comparison is further complicated because of the use of different potential energy functions.

The Thomson problem of coulombic charges on a sphere and the enumerated global minima obtained by Altschuler et al.⁵⁸ provide a quantitative test of the ability of the MDM method to find global minima among a large set of nearly degenerate states.⁵⁹ As shown in Table I, the MDM method is able to find the global minimum for systems up to 100 particles in just a few iterations, including minima not readily found using methods designed specifically for this purpose. Whereas

a lack of data concerning the efficiency of the other algorithms makes direct comparison between MDM and other methods used on this problem difficult, the results presented suggest that MDM is an efficient and accurate method for the search of potential energy surfaces with a large number of interrelated interactions.

To directly evaluate the efficiency of the molecular dynamics–minimization method proposed in this study, we choose the search of Lennard–Jones clusters as a benchmark of comparison among the different conformational search methods. This problem was chosen because the reduced Lennard–Jones potential [eq. (12)] is the same as used by many conformational search approaches. Additionally, the conformational space of Lennard–Jones clusters is more sophisticated than the one in the Thomson problem, posing a significant challenge to the different conformational search methods.

Many different conformational search methods have been applied to the problem of finding lowest energy structures of Lennard–Jones clusters. However comparison of the relative efficiency of these methods is still difficult, because of the lack of published timing information. We will concentrate on the comparison of four conformational search methods for which relative computational timing information can be deduced: the diffusion equation method (DEM⁴⁰); the method of minimization of random initial configurations by Bytheway and Kepert⁴¹ (abbreviated as BK); MDM and MCM. The Monte Carlo–minimization method was proposed previously in the context of finding lowest energy conformations of peptides⁴ and, given the analogous nature of the MDM method, we include the MCM results on Lennard–Jones clusters here for comparison. A genetic algorithm⁴⁴ has also been applied to the Lennard–Jones cluster of size $N = 13$. However, because of only one such data point, and because of lack of timing information, we have not included it in our comparison study. The simulated annealing method³⁹ has previously been applied to study Lennard–Jones clusters of up to $N = 25$; the Gaussian phase packet/Gaussian density annealing approach^{42,43} was used to study Lennard–Jones clusters of up to $N = 19$. These methods are also not included in our discussion because of lack of timing information.

The four generic conformational search methods to be compared (DEM,⁴⁰ BK,⁴¹ MDM, and MCM⁴) all share a common feature: they make use of local minimizations during the conformational search. A quantitative comparison of these four methods is

TABLE IV.
Comparison of Different Conformational Search
Methods: Total Number of Minimizations Needed
to Reach Global Energy Minimum of
Lennard–Jones Clusters.

| <i>N</i> | DEM ^a | BK ^b | MCM ^c | MDM ^d |
|----------|------------------|-----------------|------------------|------------------|
| 15 | 100–400 | 100(4) | 10–30 | 10–40 |
| 16 | — | 100(3) | 10–30 | 10–40 |
| 17 | — | 500(1) | 10–30 | 10–40 |
| 18 | — | 250(1) | 10–70 | 60–180 |
| 19 | 100–400 | 500(2) | 10–70 | 60–180 |
| 20 | — | 200(3) | 10–70 | 60–180 |
| 21 | — | 250(2) | 20–70 | 60–180 |
| 22 | — | 200(1) | 30–70 | 60–180 |
| 23 | — | 500(1) | 30–70 | 60–180 |

^a Diffusion equation method.⁴⁰ Note, for *N* = 19, a local minimum, rather than the global minimum, was found.

^b The method of minimization of random initial configurations.⁴¹ Listed in parentheses are the frequencies with which the global minimum is found among all minima obtained by minimizations of randomly generated structures.

^c The Monte Carlo–minimization method.⁴ This method was originally applied to find the global minimal energy structure for peptides. It is adapted here to study Lennard–Jones clusters for the purpose of comparison.

^d This work.

based on the total number of minimizations needed to locate the global minimum energy structures for different sizes of Lennard–Jones clusters. The results on Lennard–Jones clusters of sizes $15 \leq N \leq 23$ are listed in Table IV. On a Silicon Graphics Indigo (R4400) system, one energy minimization for a Lennard–Jones cluster of *N* = 23 takes about 4 seconds of CPU time. Thus, it takes between 0.5 and 10 minutes for the MDM method to find the global energy minima for Lennard–Jones clusters of size $15 \leq N \leq 23$.

From Table IV, it can be seen that the molecular dynamics–minimization procedure is comparatively more efficient in finding the global minimum energy structures than the diffusion equation method⁴⁰ and the method of minimizations of randomly generated structures.⁴¹ However, it is, in general, slower than the Monte Carlo–minimization method by about a factor of two. This difference in efficiency could be due to the more controlled manner in which Monte Carlo random sampling can be conducted, as compared with the “sampling” in molecular dynamics–minimization, in which a new candidate structure is generated via the velocities, which are more varied and unpredictable. As a result, relatively large propor-

tions of MDM-sampled structures are either rejected due to a failure to satisfy the energy conservation constraints [eq. (8)], or are redundant due to a failure to leave the current domain of attraction during minimization. It is not clear if the relative inefficiency of MDM sampling, as compared with MCM, is due to intrinsic limitations of the molecular dynamics formalism, or due to the particular implementation of the MDM formulation [eqs. (2)–(8)].

Although MDM is shown to be less efficient than MCM in locating the global energy minima of simple systems, such as Lennard–Jones clusters, it remains a viable alternative in the exploration of conformational space of molecular systems in a heterogeneous context (e.g., in the presence of solvent). Unlike MCM, where the conformational space is sampled by internal degrees of freedom which vary from one subsystem to another, MDM samples the conformational space using the velocities, which reflect, to varying degrees, all components of the system simultaneously. MDM explores the conformational surface incrementally in three-dimensional space, thus avoiding the problem encountered in Monte Carlo simulation of a polypeptide in solvent where a small change in the backbone torsion often leads to severe steric overlap between the polypeptide and solvent molecules. The velocities used in MDM effectively provide a “guiding potential” that increases the efficiency of conformational sampling in more complicated systems. Given these advantages, we are confident that molecular dynamics–minimization will prove to be a valuable tool in the study of conformational properties of explicitly solvated polypeptides, and such work is currently in progress.

References

1. For reviews of the multiple-minima problem and conformational search methods in the chemistry and biology literature, see: K. D. Gibson and H. A. Scheraga, in *Structure and Expression. Vol. 1: From Proteins to Ribosomes*, Academic Press, New York, 1988, p. 67; A. E. Howard and P. A. Kollman, *J. Med. Chem.*, **31**, 1669 (1988); L. Piela, K. A. Olszewski, and J. Pillardy, *J. Mol. Struct. (Theochem)* **114**, 308, 229 (1994).
2. For recent reviews of the mathematical aspects of global optimization methods, see: J. S. Arora, O. A. Elwakeil, and A. I. Chahande, *Struct. Optimization*, **9** 137 (1995); R. Horst and P. M. Pardalos, Eds., *Handbook of Global Optimization*, Kluwer, Dordrecht, 1995.

3. M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., San Francisco, 1979.
4. Z. Li and H. A. Scheraga, *Proc. Natl. Acad. Sci.*, **84**, 6611 (1987); *J. Mol. Struct. (Theochem)*, **170**, 333 (1988).
5. R. Unger and J. Moulton, *Bull. Mathemat. Biol.*, **55**, 1183 (1993).
6. M. Saunders, *J. Am. Chem. Soc.*, **109**, 3150 (1987); M. Saunders and H. A. Jimenez-Vazquez, *J. Comput. Chem.*, **14**, 330 (1993).
7. M. Lipton and W. C. Still, *J. Comput. Chem.*, **9**, 343 (1988).
8. S. R. Wilson, W. Cui, J. W. Moscovitz, and K. E. Schmidt, *Tetrahed. Lett.*, **29**, 4373 (1988).
9. G. Chang, W. Guida, and W. C. Still, *J. Am. Chem. Soc.*, **111**, 4379 (1989).
10. R. A. Dammkoehler, S. F. Karasek, E. F. B. Shands, and G. R. Marshall, *J. Comput.-Aid. Molec. Des.*, **3**, 3 (1989).
11. M. Saunders, K. N. Houk, Y.-D. Wu, W. C. Still, M. Lipton, G. Chang, and W. C. Guida, *J. Am. Chem. Soc.*, **112**, 1419 (1990).
12. J. M. Goodman and W. C. Still, *J. Comput. Chem.*, **12**, 1110 (1991).
13. H. Goto and E. Osawa, *J. Mol. Struct. (Theochem)*, **285**, 157 (1993).
14. F. Guarnieri and W. C. Still, *J. Comput. Chem.*, **15**, 1302 (1994).
15. R. Bruccoleri and M. Karplus, *Biopolymers*, **26**, 137 (1987).
16. D. R. Ripoll and H. A. Scheraga, *Biopolymers*, **27**, 1283 (1988).
17. D. S. Goodsell and A. J. Olson, *Proteins*, **8**, 195 (1990).
18. S. R. Wilson and W. Cui, *Biopolymers*, **29**, 225 (1990).
19. R. E. Bruccoleri and M. Karplus, *Biopolymers*, **29**, 1847 (1990).
20. J. J. Perez, H. O. Villar, and G. H. Loew, *J. Comput.-Aided Molec. Des.*, **6**, 175 (1992).
21. R. Abagyan and P. Argos, *J. Mol. Biol.*, **225**, 519 (1992); R. Abagyan and M. Totrov, *J. Molec. Biol.*, **235**, 983 (1994); R. Abagyan, M. Totrov, and D. Kuznetsov, *J. Comput. Chem.*, **15**, 488 (1994).
22. R. A. Abagyan, *FEBS Lett.*, **325**, 17 (1993).
23. P. Tuffery, C. Etchebest, S. Hazout, and R. Lavery, *J. Comput. Chem.*, **14**, 790 (1993).
24. U. H. E. Hansmann and Y. Okamoto, *J. Comput. Chem.*, **14**, 1333 (1993).
25. H. Meirovitch, E. Meirovitch, A. G. Michel, and M. Vasquez, *J. Phys. Chem.*, **98**, 6241 (1994); M. Vasquez, E. Meirovitch, and H. Meirovitch, *J. Phys. Chem.*, **98**, 9380 (1994).
26. T. Montcalm, W. Cui, H. Zhao, A. G. Guarnieri, and S. R. Wilson, *J. Mol. Struct. (Theochem)*, **308**, 37 (1994).
27. M. Zaslavskiy, B. A. Luty, M. E. David, and J. A. McCammon, *J. Mol. Biol.*, **238**, 455 (1994).
28. S. M. Legrand and K. M. Merz, *Mol. Simul.*, **13**, 299 (1994).
29. D. Byrne, J. Li, E. Platt, B. Robson, and P. Weiner, *J. Comput.-Aid. Molec. Design*, **8**, 67 (1994).
30. S. S. Sung, *Biophys. J.*, **66**, 1796 (1994); **68**, 826 (1994); *J. Theoret. Biol.*, **173**, 389 (1995).
31. M. Vasquez, *Biopolymers*, **36**, 53 (1995).
32. R. Srinivasan and G. D. Rose, *Proteins*, **22**, 81 (1995).
33. S. Sun, P. D. Thomas, and K. Dill, *Prot. Eng.*, **8**, 769 (1995); S. Sun, *Biophys. J.*, **69**, 340 (1995).
34. J. Hempel, R. M. Fine, M. Hassan, W. Ghoul, A. Guaragna, S. C. Koerber, Z. Li, and A. T. Hagler, *Biopolymers*, **36**, 283 (1995).
35. M. R. Hoare, *Adv. Chem. Phys.*, **40**, 49 (1979).
36. J. Farges, M. F. deFeraudy, B. Raoult, and G. Torchet, *J. Phys. (Paris)*, **38**, C247 (1977); J. Farges, M. F. deFeraudy, B. Raoult, and G. Torchet, *J. Chem. Phys.*, **78**, 5067 (1983); J. Farges, M. F. deFeraudy, B. Raoult, and G. Torchet, *Surf. Sci.*, **156**, 370 (1985); J. Farges, M. F. deFeraudy, B. Raoult, and G. Torchet, *J. Chem. Phys.*, **84**, 3491 (1986).
37. D. L. Freeman and J. D. Doll, *J. Chem. Phys.*, **82**, 462 (1985).
38. J. A. Northby, *J. Chem. Phys.*, **87**, 6166 (1987).
39. L. T. Wille, *Chem. Phys. Lett.*, **133**, 405 (1987).
40. J. Kostrowicki, L. Piela, B. J. Cherayil, and H. A. Scheraga, *J. Phys. Chem.*, **95**, 4113 (1991).
41. I. Bytheway and D. L. Kepert, *J. Math. Chem.*, **9**, 161 (1992).
42. J. Ma, D. Hsu, and J. E. Straub, *J. Chem. Phys.*, **99**, 4024 (1993).
43. C. Tsao and C. L. Brooks III, *J. Chem. Phys.*, **101**, 6405 (1994).
44. J. Mestres, and G. E. Scuseria, *J. Comput. Chem.*, **16**, 729 (1995).
45. (a) K. S. Kim, M. Dupuis, G. C. Lie, and E. Clementi, *Chem. Phys. Lett.*, **131**, 451 (1986); (b) B. J. Mhin, H. S. Kim, C. W. Yoon, and K. S. Kim, *Chem. Phys. Lett.*, **176**, 41 (1991).
46. (a) C. J. Tsai and K. D. Jordan, *J. Chem. Phys.*, **95**, 3850 (1991); (b) C. J. Tsai and K. D. Jordan, *Chem. Phys. Lett.*, **213**, 181 (1993).
47. R. J. Wawak, M. M. Wimmer, and H. A. Scheraga, *J. Phys. Chem.*, **96**, 5138 (1992).
48. K. A. Franken, M. Jalaie, and C. E. Dykstra, *Chem. Phys. Lett.*, **198**, 59 (1992).
49. A. Verairi and S. C. Farantos, *J. Chem. Phys.*, **98**, 4059 (1993); S. C. Farantos, *Z. Phys. D*, **31**, 213 (1994); B. R. Eggen, A. J. Marks, J. N. Murrell, and S. C. Farantos, *Chem. Phys. Lett.*, **219**, 247 (1994).
50. J. N. Gregory and D. C. Clary, *J. Chem. Phys.*, **102**, 7817 (1995); *Chem. Phys. Lett.*, **228**, 547 (1995).
51. H. C. Andersen, *J. Chem. Phys.*, **72**, 2384 (1980).
52. E. Bonomi, *J. Stat. Phys.*, **39**, 167 (1985).
53. L. L. Whyte, *Am. Math. Monthly*, **59**, 606 (1952).
54. L. T. Wille, *Nature (London)*, **324**, 46 (1986).
55. T. Erber and G. M. Hockney, *J. Phys. A*, **24**, L1369 (1991).
56. J. R. Edmunsen, *Acta Crystallogr.*, **A48**, 60 (1992); **A49**, 648 (1993).
57. L. Glasser and A. G. Avery, *J. Phys.*, **25**, 2473 (1992).
58. E. L. Altschuler, T. J. Williams, E. R. Ratner, F. Dowla, and F. Wooten, *Phys. Rev. Lett.*, **72**, 2671 (1994).
59. T. Erber and G. M. Hockney, *Phys. Rev. Lett.*, **74**, 1482 (1995).
60. E. L. Altschuler, T. J. Williams, E. R. Ratner, F. Dowla, and F. Wooten, *Phys. Rev. Lett.*, **74**, 1483 (1995).
61. F. H. Stillinger and D. K. Stillinger, *J. Chem. Phys.*, **93**, 6013 (1990).
62. P. A. Braier, R. S. Berry, and D. J. Wales, *J. Chem. Phys.*, **93**, 8745 (1990).

63. R. S. Berry, *J. Phys. Chem.*, **98**, 6910 (1994).
64. (a) K. Liu, J. D. Cruzan, and R. J. Saykally, *Science*, **271**, 929 (1996); (b) J. D. Crusan, L. B. Braly, K. Liu, M. G. Brown, J. G. Loeser, and R. J. Saykally, *Science*, **271**, 59 (1996); (c) K. Liu, M. G. Brown, J. D. Cruzan, and R. J. Saykally, *Science*, **271**, 62 (1996); (d) K. Liu, M. G. Brown, C. Carter, R. J. Saykally, J. K. Gregory, and D. C. Clary, *Nature*, **381**, 501 (1996).
65. M. F. Vernon, D. J. Krajnovich, H. S. Kwak, J. M. Lisy, Y. R. Shen, and Y. T. Lee, *J. Chem. Phys.*, **77**, 47 (1982); R. H. Page, J. G. Frey, Y. R. Shen, and Y. T. Lee, *Chem. Phys. Lett.*, **106**, 373 (1984).
66. J. C. Owicki, L. L. Shipman, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 1794 (1975); A. W. Burgess, L. L. Shipman, and H. A. Scheraga, *J. Am. Chem. Soc.*, **98**, 23 (1978).
67. (a) G. Brink and I. Glasser, *J. Phys. Chem.*, **88**, 3412 (1984); (b) F. H. Stillinger and C. W. David, *J. Chem. Phys.*, **73**, 3384 (1980).
68. R. O. Watts, *Chem. Phys.*, **26**, 367 (1977).
69. (a) O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.*, **64**, 1351 (1974); (b) G. C. Lie, E. Clementi, and M. Yoshimine, *J. Chem. Phys.*, **64**, 2314 (1974); (c) D. A. Estrin, L. Paglieri, G. Corongiu, and E. Clementi, *J. Phys. Chem.*, **100**, 8701 (1996).
70. P. Cieplak, P. Kollman, and T. Lybrand, *J. Chem. Phys.*, **92**, 6755 (1990).
71. C. E. Dykstra, *J. Comput. Chem.*, **9**, 476 (1988); *J. Am. Chem. Soc.*, **111**, 6168 (1989).
72. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **79**, 926 (1983); W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988); W. L. Jorgensen and J. M. Briggs, *J. Am. Chem. Soc.*, **111**, 4190 (1989).
73. M. Levitt, M. Hirshberg, R. Sharon, and V. Daggett, *Comput. Phys. Comm.*, **91**, 215 (1995); M. Levitt, M. Hirshberg, R. Sharon, K. E. Laidig, and V. Daggett, *J. Phys. Chem.*, **101**, 5051 (1997).
74. J. E. Fowler and H. F. Schaefer, *J. Am. Chem. Soc.*, **117**, 446 (1995).
75. J. Pillardy and L. Piela, *J. Phys. Chem.*, **99**, 11805 (1995). These investigators report finding a lower energy minimum for $N = 38$ than Northby.³⁸